**Minutes for the Geuvadis RNAseq TC from May 6 2011**

Written by Thomas Giger

University of Geneva

**1. As of now we made plans that the different labs analyze samples from one particular population (as listed in the workflow document that I've sent around on April 7).**

**It would be a better experimental design if the samples were randomly distributed among the different labs. Since there is still time to react - do we want to switch the strategy?**

We agreed that every lab will analyze a random set of samples instead of individuals coming from only one population. We believe that the technology is robust enough that we can do this, even though that there may be factors that we can’t control in reality (e.g. people will use slightly different kinds of equipment like different kinds of incubators etc.) which may lead to a better comparability of data produced within the same lab than between different labs. But one of the main goals of our collaborative effort is to learn and demonstrate how to do such a project involving distributed sequencing – since future experiments will comprise an order of magnitude more samples and distributed sequencing will become more and more important. Towards this aim we decided that every lab will sequence - in addition to the number of samples they committed to earlier – 3 to 5 samples that will be analysed in every lab. Those samples will help to normalize data across the different labs. For the same purpose the lab in Geneva will sequence a 10% overlap of the samples that are analysed in every other lab.

Also during the experiment people are requested to keep track on the equipment they use (e.g. kind of incubator, provider etc.) and share this information with the data analysis group.

**2. How long are the library prep / sequencer queue up times in the different labs (once you receive the RNA samples - how long do you think will it take you to analyze them?).**

**If some labs expect to have much larger delays, we could make arrangements that those labs get samples earlier than other labs.**

**This would also mean that we would all use the protocol that is used at the time point when the first lab starts to analyze samples.**

During the call it became apparent, that the queuing times for a sample in the wet lab analysis pipeline are about the same in every lab - around 3 months. We therefore decided that RNA samples will be distributed all at the same time – which is expected to be in the beginning of September. About a month before that the lab in Geneva and the labs in Barcelona will provide the final information on the protocols to be used. The point here is that every lab orders chemicals around the same time to make sure that we don’t run into version problems of the illumina chemistry anymore. This also means that the data set is expected to be generated around December 2011.

**3. Which other samples (other than the 500 1KGP European samples) should be analyzed by RNAseq?**

This point remains to be further discussed with Stefan Schreiber and other people that are working on the exome sequencing of the disease samples.

**4. Specifications for the small RNA protocol.**

Final protocol parameters will be provided one month before the samples are distributed - Peter suggested that we should do 36 cycles single end sequencing for at least 3 million reads. We agreed to go ahead with this startegy.

Additional remarks on protocols : For the sequencing of messenger RNAs we should aim for 15-20 million reads that map to coding sequence per sample (note this is different from the 7-10 million clusters that we were talking about at earlier timepoints). How many samples people pool (using indexing) in a particular run has to be decided by every lab itself.

Finally I need the following information from you :

* Please provide me with contact information of a person in your lab that is responsible for receiving the frozen RNA samples.

- Will you be sequencing the samples on a Hiseq and/or GAII ?

* Please make sure that you will have the space in your liquid nitrogen tank / or -150 °C freezer for the 500 cell line samples (5 boxes containing 100 cryovials (1.8 ml) each) by mid September

Kindest regards,

Thomas Giger